Host-Pathogen Interaction Session

Bioinformatics Resource Center (BRC) for Biodefense and Emerging/Re-emerging Infectious Diseases Kick-Off Meeting - 13OCT2004

Switch Gears

- ***Start to think beyond the organism in isolation***
- ***Start to think beyond the genomic sequence data***
- ***Start to think about experimental problems from the viewpoint of the user (use cases)***

Goal of Session

Assuming we agree that BRCs should be a combination of informational web site/research database/data mining & analytical resource

•To begin to define the kinds of data that contribute to an understanding of the host/pathogen interaction

•To begin to define the kinds of analytical and data mining approaches that are useful in understanding the host/pathogen interaction

•To begin to define the computational structure that should be used to support research aimed at understanding the host/pathogen interaction

•To identify BRC representatives who would form a working group (HPWG) for the development of a computational framework that can support host/pathogen interaction data during the life of the BRC contract

Approach

- Identify key concepts/problems relevant to this research community

- Define experiments used to address the key concepts/problems

- Define components of experiments

- Describe the analytical process used to infer knowledge from the experiments

PopGen Program

- Study at McMaster University by Mark Loeb

- Goal is to map genes related to disease severity in patients infected with West Nile Virus in US

- Specific Aim is to analyze genetic polymorphisms (SNPs) in a large number of immune-related genes (HLA, cytokines, chemokines, signaling proteins, transcription factors, etc.)

- Sample set - 1200 cases with encephalitis, meningitis, acute flaccid paralysis & 1200 cases who were infected but did not develop neuroinvasive disease

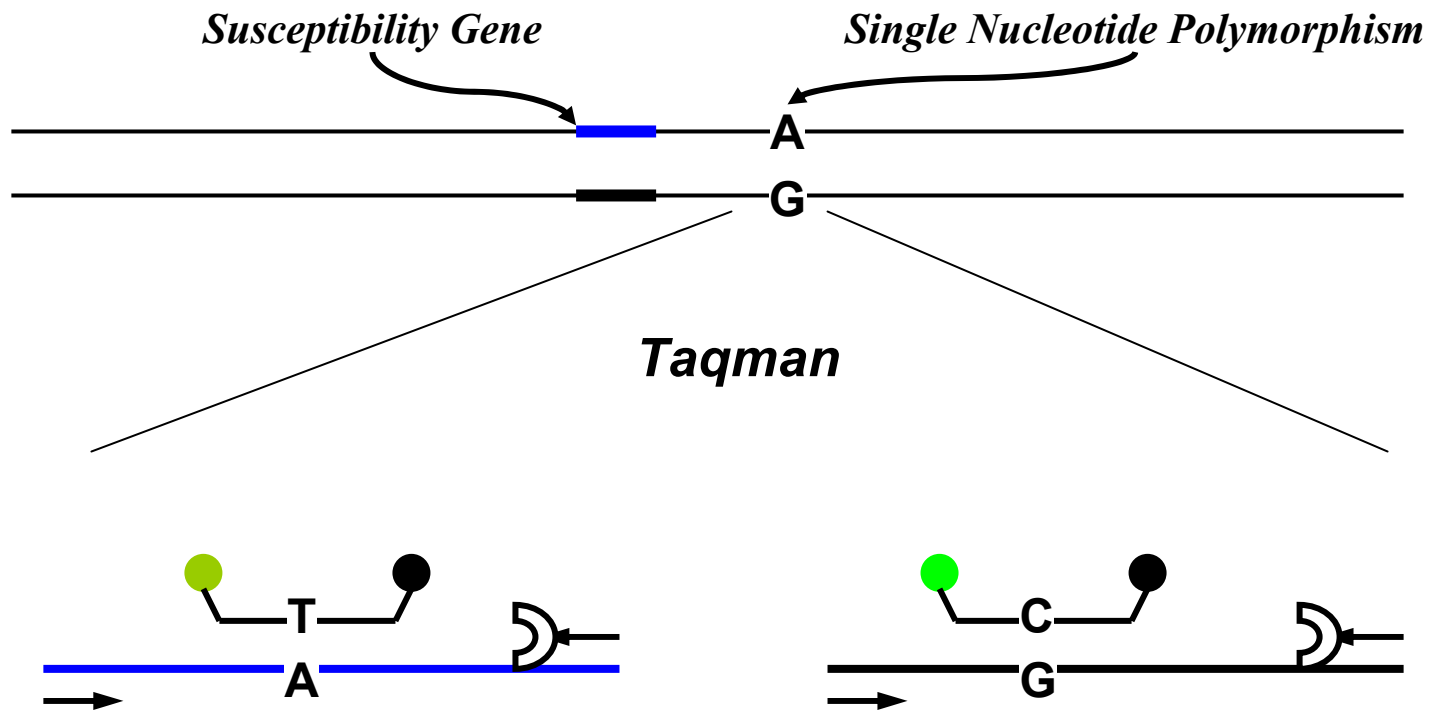- Example of a study to define *host susceptibility*

Define use case

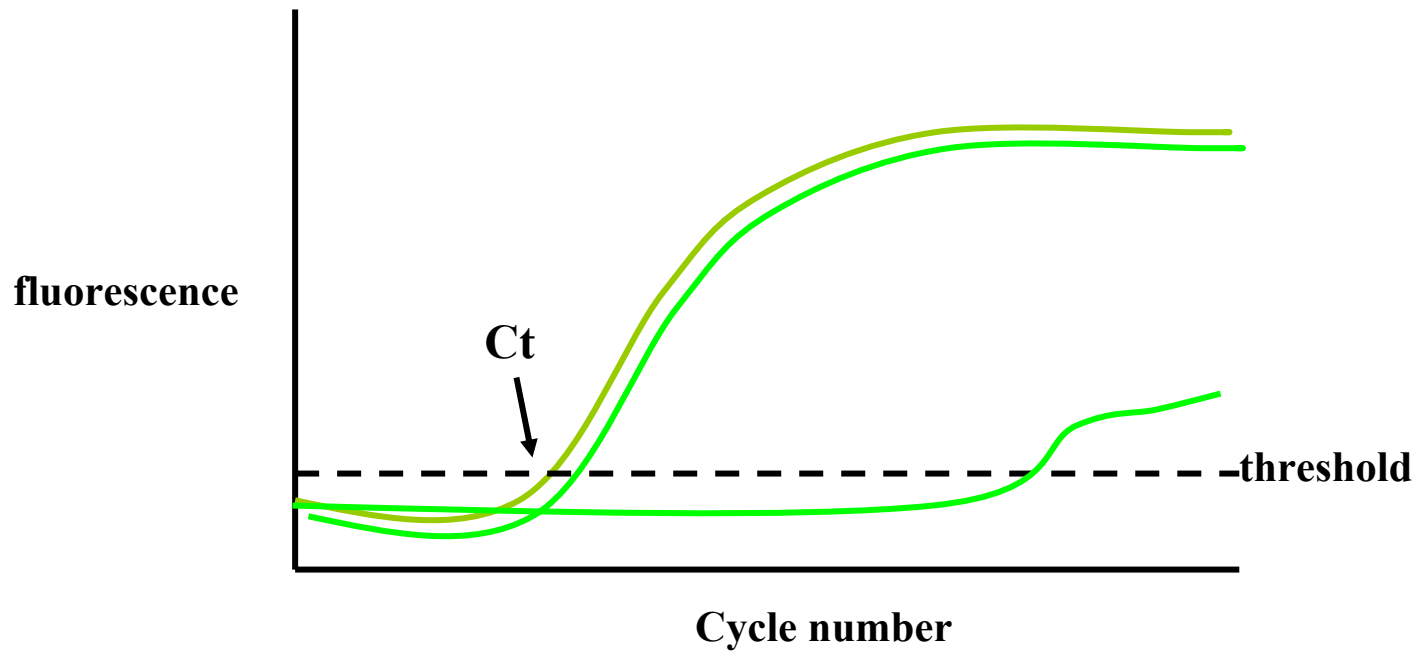[Use case example.doc](Use%20case%20example.doc)

Data to support experimental design

- dbSNP
- Clinical data
  - Patient demographics
  - Patient medical history
  - Physical exam
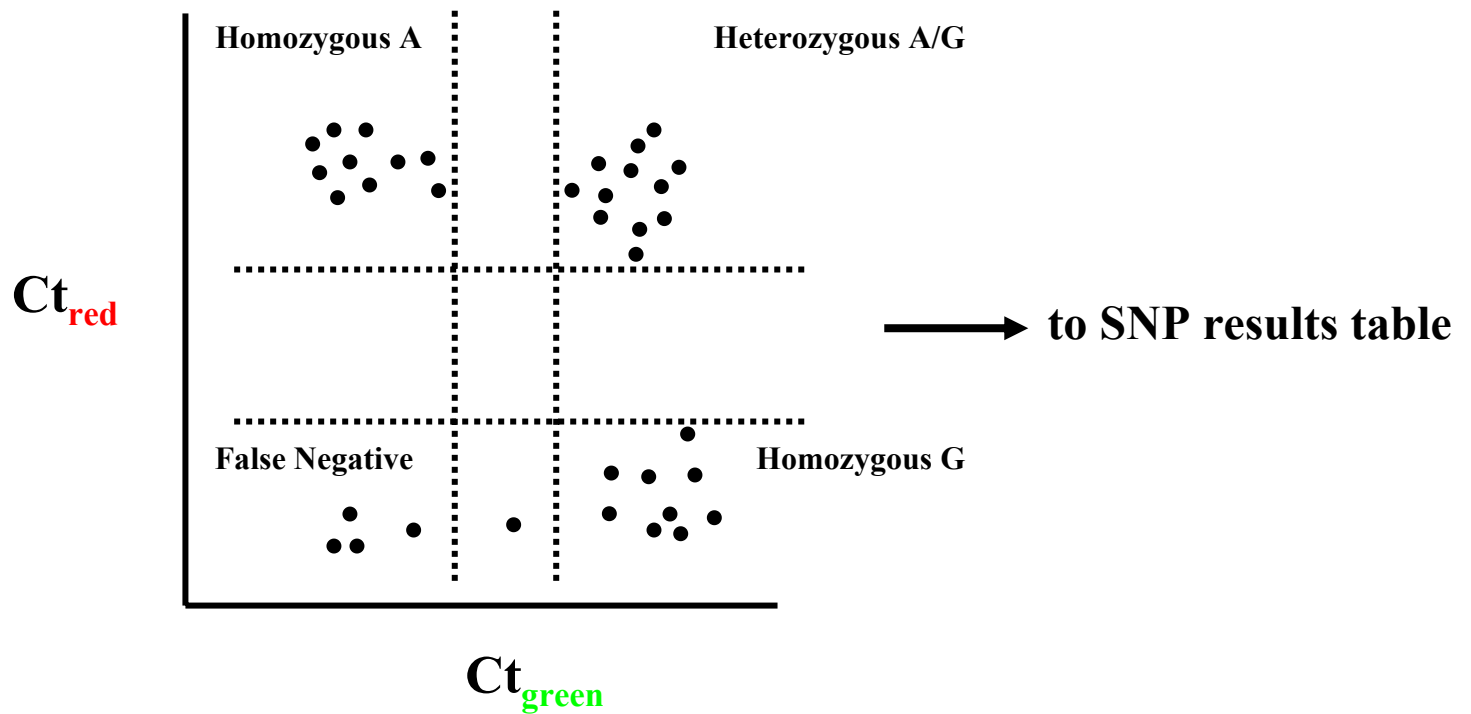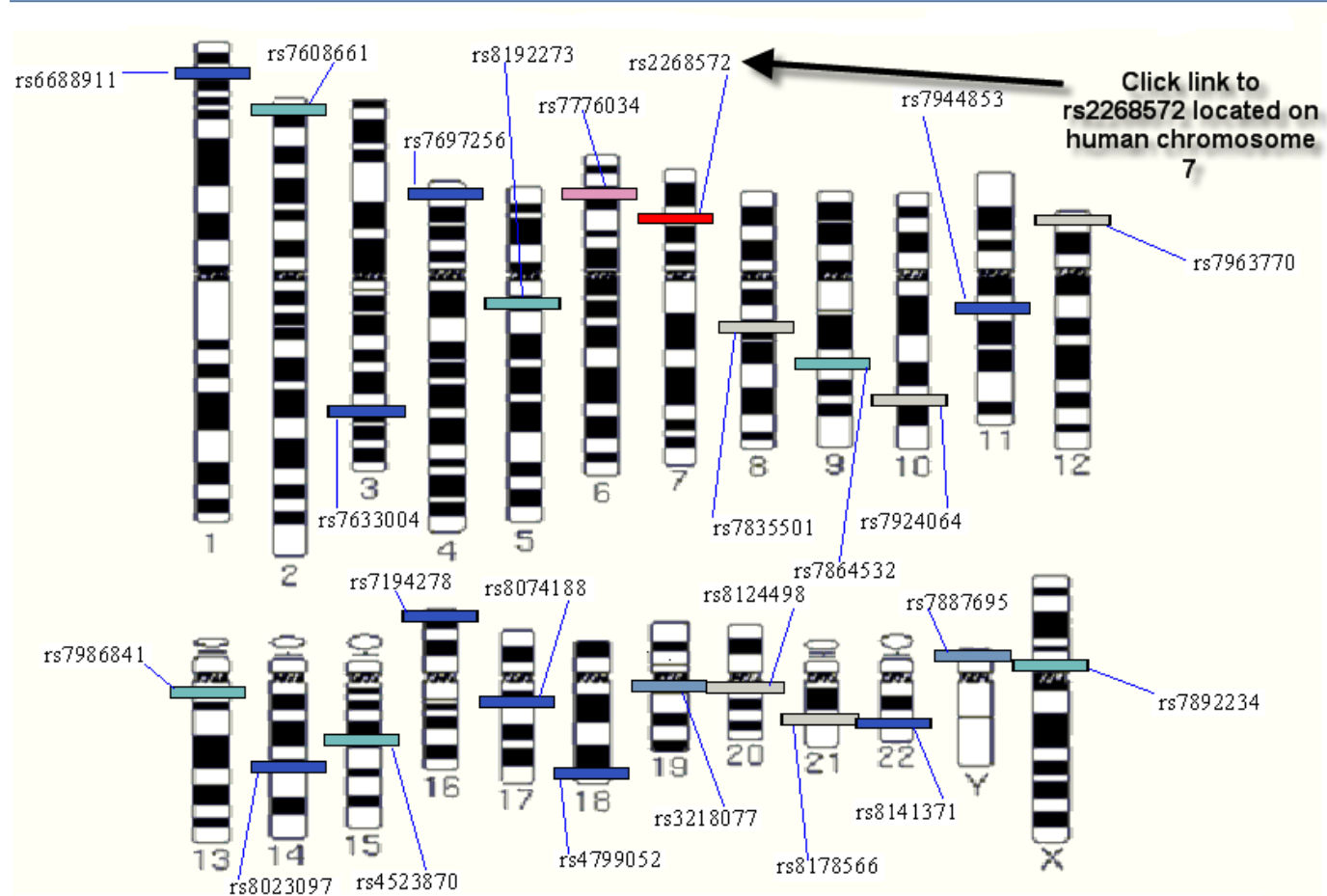 - Diagnostic laboratory tests

Genetic markers (SNP)

Susceptibility Gene

Single Nucleotide Polymorphism

A

G

Taqman

T   A

C   G

Realtime PCR raw data

Data Processing Tool

SNP Haplotype Table
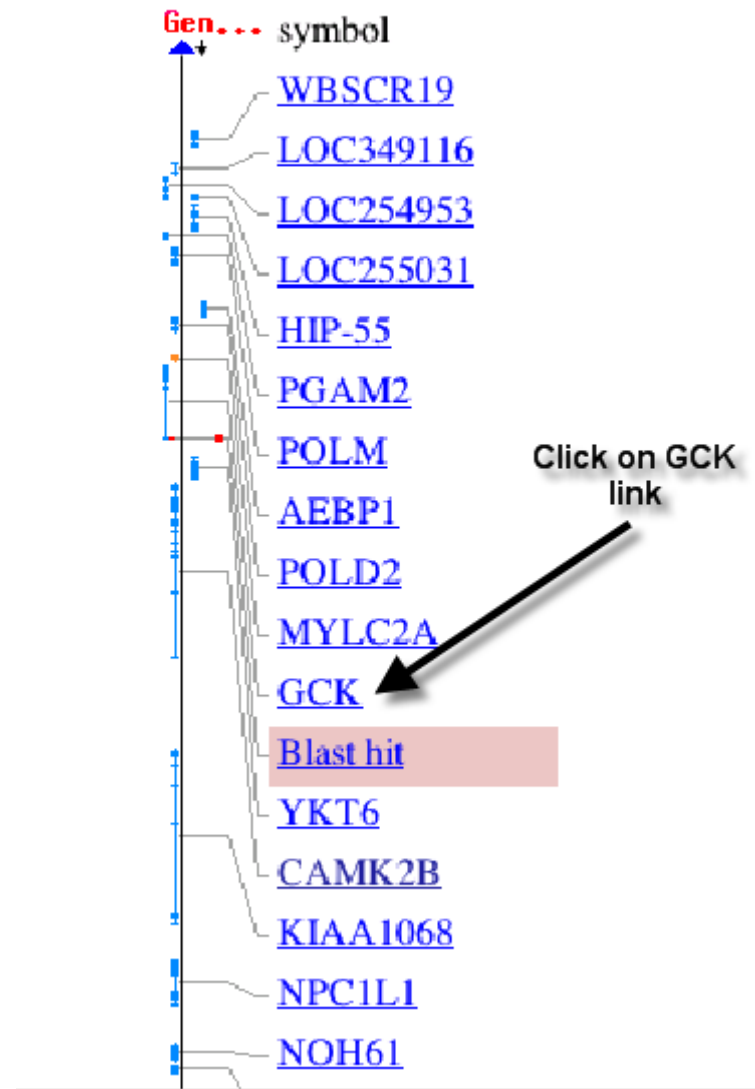
SNP haplotype table.xls

Chromosome view

High resolution view

High Resolution Physical Map of Human Chromosome 7

Summary

- Host-pathogen concepts
  - Disease susceptibility
  - Clinical manifestations
- Data types
  - Experimental design
    - Gene annotation
    - SNP information
    - HLA information
    - Clinical information
  - Experimental data
    - Raw data (Ct values)
    - Processed data (Chi-squared p value)
- Data analysis
  - PCR data processing
  - Association statistics
- Data interpretation
  - Disease association and linked genes

Identification of the key concepts relevant to the host-pathogen interactions

- Virulence Determinants
  - Genes
  - Mechanisms of action
- Host Susceptibility
  - Genes and their functions
  - Non-genetic factors
- Combination genetics - host susceptibility combined with pathogen virulence (donor-recipient tissue typing paradigm)
- Host/pathogen factor interactions
- Ontology development (know when to stop!)
  - From the perspective of the experiments
  - Things you measure vs interpretation
- Immune Response
- Pathogenic Epitopes
- Vaccine Development
- Role of vector
- Disease Pathogenesis/Host response
  - Clinical manifestations of infection
  - Physiological manifestations of infection
  - Cellular manifestations of infection
  - Molecular manifestations of infection
- Transmission
- Epidemiology (MIDAS)
- Therapeutic Intervention
- Drug Resistance
- Diagnostic targets

Session

- A. With the key concepts defined, what kinds of experimental designs, approaches, and methodologies are used to investigate these concepts?

- B. Can these various experimental approaches be described using a common data structure?

- C. What is the nature of the results data generated from these experiments?

- D. Can the experimental results be described using a common data structure?

- E. What is the nature of the data analysis applied to the experimental results?

- F. Can the analytical process be described using a common data structure?

- G. What is the nature of the biological knowledge inferred from the analysis of the experimental data?

- H. Can the inferred knowledge be described using a common data structure?

Ontology

- A.   Discuss the importance of using a common ontology for the BRC project

- B.   Discuss the relationship between the ontology and the data model

- C. Discuss if an existing ontology would meet the needs related to host/pathogen interactions, e.g. GeneOntology, NCI Thesaurus, SNOMED, etc.